

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 15:50:54 ; Search time 2481.53 Seconds
(without alignments)
16465.781 Million cell updates/sec

Title: US-09-497-967-5
Perfect score: 1404
Sequence: 1 atgaagaacacatctctggt.....tgatctcttactacctgtg 1404
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.sts.*
27: em.un.*
28: em.vi.*
29: em.htg_hum.*
30: em.htg_inv.*
31: em.htg_inv.*
32: em.htg_inv.*
33: em.htg_inv.*
34: em.htg_pln.*
35: em.htg_pln.*
36: em.htg_pln.*
37: em.htg_vrt.*
38: em.htg_vrt.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	782.6	55.7	3026	3	AF324424	AF324424 Ichthyoph
2	63.2	4.5	2486	3	AF140273	AF140273 Ichthyoph
3	55.2	3.9	125020	9	AF429315	AF429315 Homo sapi
4	54.4	3.9	1520	3	AF405431	AF405431 Ichthyoph
5	52.2	3.7	1249	3	ICYMANT	M92907 Ichthyophth
6	50	3.6	125020	9	AF429315	AF429315 Homo sapi
7	43.4	3.1	230372	2	AC073693	AC073693 Mus muscu
8	41.4	2.9	159764	9	AC093866	AC093866 Homo sapi
9	41	2.9	643	8	AF374493	AF374493 Thalassio
10	41	2.9	179217	2	AC099748	AC099748 Bos tauru
11	39.4	2.8	591	8	AF374552	AF374552 Thalassio
12	38.6	2.7	643	8	AF374492	AF374492 Thalassio
13	38.6	2.7	16814	9	HSCOLPDGF	X98705 H.sapiens D
14	38.6	2.7	341950	1	AP003600	AP003600 Nostoc sp
15	38.4	2.7	224271	9	AC026803	AC026803 Homo sapi
16	38.2	2.7	15294	1	AE006936	AE006936 Mycobacte
17	38.2	2.7	31300	1	MTV035	AL021930 Mycobacte
18	38.2	2.7	152777	2	AC125855	AC125855 Rattus no
19	38.2	2.7	233345	4	AJ421481	AJ421481 Bos tauru
20	38.2	2.7	338150	1	AP005278	AP005278 Corynebac
21	38.2	2.7	349980	6	AX127147	AX127147 Sequence
22	37.8	2.7	559	8	AF374543	AF374543 Thalassio
23	37.8	2.7	591	8	AF374540	AF374540 Thalassio
24	37.8	2.7	591	8	AF374542	AF374542 Thalassio
25	37.8	2.7	591	8	AF374544	AF374544 Thalassio
26	37.8	2.7	591	8	AF374545	AF374545 Thalassio
27	37.8	2.7	591	8	AF374547	AF374547 Thalassio
28	37.8	2.7	591	8	AF374548	AF374548 Thalassio
29	37.8	2.7	591	8	AF374549	AF374549 Thalassio
30	37.8	2.7	591	8	AF374550	AF374550 Thalassio
31	37.8	2.7	591	8	AF374551	AF374551 Thalassio
32	37.8	2.7	643	8	AF374490	AF374490 Thalassio
33	37.8	2.7	643	8	AF374491	AF374491 Thalassio
34	37.8	2.7	643	8	AF374492	AF374492 Thalassio
35	37.8	2.7	643	8	AF374493	AF374493 Thalassio
36	37.8	2.7	643	8	AF374494	AF374494 Thalassio
37	37.8	2.7	643	8	AF374495	AF374495 Thalassio
38	37.8	2.7	643	8	AF374496	AF374496 Thalassio
39	37.8	2.7	643	8	AF374497	AF374497 Thalassio
40	37.8	2.7	643	8	AF374498	AF374498 Thalassio
41	37.8	2.7	643	8	AF374499	AF374499 Thalassio
42	37.8	2.7	643	8	AF374500	AF374500 Thalassio
43	37.8	2.7	643	8	AF374501	AF374501 Thalassio
44	37.8	2.7	643	8	AF374502	AF374502 Thalassio
45	37.8	2.7	643	8	AF374503	AF374503 Thalassio
						AF374504 Thalassio
						AF374505 Thalassio

ALIGNMENTS

RESULT 1
AF324424
LOCUS
DEFINITION
AF324424
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AF324424
Ichthyophthirius multifiliis
(IAG52A) gene, complete cds.
AF324424
GI:12698726
Ichthyophthirius multifiliis.
Ichthyophthirius multifiliis
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
Hymenostomatida; Ophryoglenina; Ichthyophthirius.
1 (bases 1 to 3026)
Lin, Y., Lin, T.L., Wang, C.C., Wang, X., Stieger, K., Klopfleisch, R.
and Clark, T.G.

AF324424
3026 bp
DNA
linear
INV 27-FEB-2002
Ichthyophthirius multifiliis
immobilization antigen isoform

TITLE Variation in primary sequence and tandem repeat copy number among i-antigens of Ichthyophthirius multifiliis
JOURNAL Mol. Biochem. Parasitol. 120 (1), 93-106 (2002)
MEDLINE 21839613
PUBMED 11849709
REFERENCE 2 (bases 1 to 3026)
AUTHORS Lin, Y., Lin, T.-L. and Clark, T. G.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-2000) Microbiology and Immunology, NYSCVM Cornell University, Ithaca, NY 14853, USA
FEATURES
source Location/Qualifiers
1..3026
/organism="Ichthyophthirius multifiliis"
/macronuclear
/serotype="D"
/isolate="C5"
/db_xref="taxon:5932"
1285..2817
/gene="IAG52A"
1285..2817
/gene="IAG52A"
/product="immobilization antigen isoform"
1330..2736
/gene="IAG52A"
/note="52 kDa protein"
/codon_start=1
/transl_table=6
/product="immobilization antigen isoform"
/protein_id="AAK01661.1"
/db_xref="GI:12698727"
/translation="MKNILVILISLFINOIKSANCPVGTETINTAGQVDDLGT PAN C VNCQNFYNNAAAFVPGASTCTPCPKKDGADQNPATANLVTCNVKCPAGTATA GADTAAAIITECVNCIRNFYNENAPNFNAGSTCTACPNRVGGALTAGNAATIVAG CNVACPTGLDDGVTDYVRSFTEVKCRUNFYNNNGNTPFPNPKSQCTPCPAIK PANVAQATILGNDAITIAQCNVACPDGTISAAGVNNWVAQNTCEINCAFPNFYNNAPNF NPNSTNCLDPKANDYGAEATAGGAATLAKOCNIACPDGTIAAGSANTYVILQTECLN CAANFYEDCPNKGSSRCKACPNKVOGAVATAGGTATLIAQCALECPAGCTVLDGT TSYKQAASECVKCAANFYTTQTQDWAGIDTCTSCNKKLTSGAEANLPESAKKNIQC DFANFLISILSLLISLYLL"

source

gene

mRNA

CDS

BASE COUNT 1177 a 316 c 355 g 1178 t
ORIGIN

Query Match 55.7%; Score 782.6; DB 3; Length 3026;
Best Local Similarity 72.5%; Pred. No. 8.3e-218;
Matches 1013; Conservative 0; Mismatches 384; Indels 0; Gaps 0;

QY 1 ATGAAGAACAAATCTGGTGATCCTGATCATCTCTCTGTTTCATCAACAGATCAAGTCT 60
DB 1330 ATGAAAAATAATATTTAGTAAATATTGATTATTTTCATTATTTATCAATTAATATCT 1389
QY 61 GCTAACTGCTGTGGAGCGAGACCAACACCGCTGGACAGGTGGACGACCTGGGAACC 120
DB 1390 GCTAATGTGCTTGGAACTGAACCTAACACACCGGATAAGTTGATGATCTAGGAAC 1449
QY 121 CTTGCTAACTGTGGAACGTGAGAAGAACTTCTACTACACACGCTGCTGCTTCGTG 180
DB 1450 CTGCAAAATGTTGTAATTTGTAGAAAAACTTTTATTATAATAATGCTGCTTTCGTT 1509
QY 181 CTGGAGCTTCTACCTGTACCCCTTGTCTCAGAGAAGGACGCTGGAGCTCAGCCCTAAC 240
DB 1510 CTGTGTGCTAGTACGTACAGCTTGTCCATTAATAAAGATGCTGGTCTTAACCAAT 1569
QY 241 CTCTCTGTACCGCTAACCTGTGTGACCCAGTGTAAACGTGGAAGTGTCTGTGGAACCGCT 300
DB 1570 CCACCTGCTACTGCTAATTTAGTCACATAATGTAACGTTAAATGCCCTGTGTACCGCA 1629
QY 301 ATCGCTGGAGAGGTACCGACTACGCTGCTATCATCCAGGCTGTGGAACCTGCGGATC 360
DB 1630 ATTGCAGGTGGAGCAACAGATTATGCACAAATATCAAGAAATGTTAATTTAGTAAT 1689
QY 361 AACTTCTACACGAGACGCTCTCACTTCAACGCTGGAGCTTCTACCTGTACCGCTGTG 420
DB 1690 AATTTTTTAATGAAATGCTCCAAATTTTAATGCAGGTGCTAGTACATGCACAGCTTGT 1749

QY 421 CCTGTGAACCGGTGGAGGAGCTCTGACCGCTGAAACGCTGCTACCATCTGGCTCAG 480
DB 1750 CCGGTAAAACAGAGTTGGTGGTGCATTTGACTGCTGTAATCCCGCTACCATAGTCGATA 1809
QY 481 TGTAACTGGCTTGTCTTACCGGAACCGCTCTGGACGACGAGTGACCCGACTAGGTG 540
DB 1810 TGTAACTGGCTTGTCTTACCGGAACCGCTCTGGATGATGAGTAAGTATGATTT 1869
QY 541 CGCTCTTTACCGAGTGTGTGAGTGTGCTGAACTTCTACTACAACGGAACAAACGGA 600
DB 1870 AGATCATTTACAGAAATGTTAAATGTAGACTTAACITTTACTATATGTTAATAATGGT 1929
QY 601 AACACCCCTTTCAACCCCTGGAAAGTCTCAGTGTACCCCTTGTCTCTGTATCAACCTGCT 660
DB 1930 AATACCTCTTCAATCCAGGTAAAGTTAATGCACACCTTGTCCGGCAATTAACCTGCT 1989
QY 661 AAGTGGCTCAGGTACCCCTGGGAACGACGCTACCATCACCCTCAGTGTACCGTGGCT 720
DB 1990 AATGTTGCTTAAGCTACTTTAGTAAATGATGCTACATAACCGCATATGTAACTGGA 2049
QY 721 TGTCTTGACGGAACCATCTCTGCTGCTGGAGTGAACACTGGGTGGCTCAGAACACGAG 780
DB 2050 TGGCTGTGATGCTACTAAGTGTCTGCTGGAGTAAATAATTTGGTAGCACAAAACACTGAA 2109
QY 781 TGTACCAACTGTGCTCCTTAACCTTCTACACACACGCTCTCAACTTCAACCTCGGAAC 840
DB 2110 TGTACTTAATTTGCTCTAATCTTACAAATAATAATGCTGCTGCTGCTGCTGCTGCTGCT 2169
QY 841 TCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2229
DB 2170 AGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2289
QY 901 GCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2349
DB 2230 GCGCTACTTTAGCAAAATTAATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2469
QY 961 GGAGCTACCAACTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 2290 GGAGCAACTAATATGATTAATTAATTAACAGAAATGCTAAATTTGCTGCTGCTGCTGCTGCT 2349
QY 1021 TTCGACGGAACAACTTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 2350 TTTGATGCTAATATTTCTAGGCGAGGAAGTAGTAGATGCAAGCATGTCCAGCAATAAA 2409
QY 1081 GTGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 2410 GTTTAAGCGCTGTAGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2469
QY 1141 GAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 2470 GAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1201 TCTGAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 2530 TCTGAATGTGTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2589
QY 1261 GAAATCGACACTGTACCTTCTTGAACAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 2590 GGTATTTGATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2649
QY 1321 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
DB 2650 CCGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2709
QY 1381 CTGCTGATCTCTTACTA 1397
DB 2710 TTATTGATTTCTTATTA 2726

RESULT 2
AF140273 AF140273 2486 bp DNA linear INV 19-MAY-1999
LOCUS

DEFINITION	Ichthyophthirius multifiliis immobilization antigen precursor (IAG48) gene, complete cds.
ACCESSION	AF140273
VERSION	AF140273.1
KEYWORDS	GI:4868370
SOURCE	Ichthyophthirius multifiliis.
ORGANISM	Ichthyophthirius multifiliis
REFERENCE	Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida; Ophryoglenina; Ichthyophthirius.
AUTHORS	1 (bases 1 to 2486) Clark,T.G., Lin,T.L., Jackwood,D.A., Sherrill,J., Lin,Y. and Dickerson,H.W.
TITLE	The gene for an abundant parasite coat protein predicts tandemly repetitive metal binding domains
JOURNAL	Gene 229 (1-2), 91-100 (1999)
MEDLINE	99196987
PUBMED	10095108
REFERENCE	2 (bases 1 to 2486) Gaertig,J., Gao,Y., Tishgarten,T., Clark,T.G. and Dickerson,H.W.
AUTHORS	Surface display of a parasite antigen in the ciliate Tetrahymena thermophila
TITLE	3 (bases 1 to 2486) Clark,T.G., Lin,T.-L., Jackwood,D.A. and Dickerson,H.W.
JOURNAL	Direct Submission
REFERENCE	Submitted (29-MAR-1999) Microbiology & Immunology, Cornell University, College of Veterinary Medicine, Ithaca, NY 14853, USA
AUTHORS	Location/Qualifiers
TITLE	1. .2486 /organism="Ichthyophthirius multifiliis" /serotype="A" /isolate="GI" /db_xref="taxon:5932" 399. .1845 /gene="IAG48" 399. .1845 /gene="IAG48" /product="immobilization antigen precursor" 433. .1761 /gene="IAG48" /note="surface protein; 48 kDa I-antigen" /codon_start=1 /trans_table=6 /product="immobilization antigen precursor" /protein_id="AAD31283.1" /db_xref="GI:4868371" /translation="MKINLIILILISFLINELRAYPCPDGTQAGLTDYGAADLGT VNCRFNYNGGAAGKANGNPFAANNAARGICVPCQINRVGSVFNAGDLATLATQ STQCPGTALDDGVTDFVRSAAQCCKPKNFYNGSPQGEAPGVQFAAGAAAG AAVTSQCVQGLKNDSPATGAQAQLATQCSNOCPTGTVLDDGTVLVENTSLVC CRPNFYNGGSPQGEAPGVQFAAGAAAGAAVTSQCVQCIKNDSPATGAQAQ ATQCSYCPGTGTAIQDGVTLVFNSTSTQCSQCIANFFNGNFAGKSQCLKCPVSK PAHAPGATNTQAQTCILCTCPAGTVLDGTSFNVASATFKCSAGFFASKTKTGT TDTCTGCTKLTSGATAKVAEATQKVQCASTTTFLSLISLFIYPILL"
FEATURES	source
gene	
mRNA	
CDS	
sig_peptide	896 a 310 c 321 g 959 t
mat_peptide	
BASE COUNT	
ORIGIN	

[illegible]

RESULT 5	Query Match	3.7%	Score 52.2	DB 3	Length 1249
ICYIMANT	Best Local Similarity	52.5%	Pred. No. 0.00089		
LOCUS	Matches 114	Conservative 0	Mismatches 103	Indels 0	Gaps 0
DEFINITION	Ichthyophthirius multifiliis immobilization antigen precursor, mRNA, partial cds.				
ACCESSION	M92907				
VERSION	M92907.1				
KEYWORDS	GI:3628568				
SOURCE	Ichthyophthirius multifiliis.				
ORGANISM	Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida; Ophryoglenina; Ichthyophthirius. 1 (bases 1 to 1249)				
REFERENCE	Lin, T.L. and Dickerson, H.W. Purification and partial characterization of immobilization antigens from Ichthyophthirius multifiliis J. Protozool. 39 (4), 457-463 (1992)				
AUTHORS	93020590				
TITLE	1383510				
JOURNAL	2 (bases 1 to 1249)				
MEDLINE	Clark, T.G., McGraw, R.A. and Dickerson, H.W. Developmental expression of surface antigen genes in the parasitic ciliate Ichthyophthirius multifiliis Proc. Natl. Acad. Sci. U.S.A. 89 (14), 6363-6367 (1992)				
PUBMED	92335298				
REFERENCE	1631132				
AUTHORS	3 (bases 1 to 1249)				
TITLE	Clark, T.				
JOURNAL	Direct Submission				
MEDLINE	Submitted (06-JUN-1992) Microbiology and Immunology, Cornell University, Ithaca, NY 14853, USA				
PUBMED	4 (bases 1 to 1249)				
REFERENCE	Clark, T.				
AUTHORS	Direct Submission				
TITLE	Submitted (18-SEP-1998) Microbiology and Immunology, Cornell University, Ithaca, NY 14853, USA				
JOURNAL	Sequence update by submitter				
COMMENT	On Sep 18, 1998 this sequence version replaced gi:159289.				
FEATURES	Location/Qualifiers				
source	1. .1249				
	/organism="Ichthyophthirius multifiliis"				
	/strain="Georgia"				
	/db_xref="taxon:5932"				
	/dev_stage="trophont"				
	<1. .1190				
	/note="48 kDa"				
	/citation=[2]				
	/citation=[1]				
	/codon_start=3				
	/evidence=experimental				
	/transl_table=6				
	/product="Immobilization antigen precursor"				
	/protein_id="AAC36158.1"				
	/db_xref="GI:3628569"				
	/translation="AVPCPDGTQTQAGLTDVGAADLGTVCNCRPNFYNGSAAQGEAN GNPFANNAARIGICVPCOINRVGSVTNAGDLATLQCSQCTGTALDDGVTDFD RSAAQCVKCKPNFYNGSGPOGEAGVQVFAAGAAAGVAAVTSQCVCQLNKNDSPA TAGAANLATQCSNCGPTGTVLDDGVTILVFNTSATLCKRPNFYNGSGPOGEAGPVQVFAAGAAAGVAAVTSQCVCQLNKNDSPATAGANLATQCSQCTGTATQDQVTV LVFNSSTQCSQCIANYEFNLEAGKSQCLKCPVSKTTPAHAPGNATQATQCLTTC PAGTVLDDGTSTNFVASATECTKCSAGFFASKTTGTAGTDTCTECTKLTSGATAVN IHQ"				
	1. .5				
	6. .1187				
	/product="immobilization antigen"				
	/evidence=experimental				
	177. .893				
	/citation=[2]				
	/rpt_type=tandem				
	/rpt_unit=414. .680				
sig_peptide	348 a				
mat_peptide	243 c				
repeat_region	257 g				
	401 t				
BASE COUNT					
ORIGIN					

QY	1098	TACCGCTGGAGGAACCGCTACCCCTGATCGCTCAGTGTCTCTGGAGTGTCTCGTCTGCTGAAC	1157	
Db	956	TGCTCCAGGTAACTACTTAAAGCCACATAAATGTTGACCATGTCTCTCGTCTGCTGAAC	1015	
QY	1158	CGTCTGACCGAGCAACCTTACCTACACGAGCTGCTTCTGAGTGTGTGAAGTGC	1217	
Db	1016	AGTACTTGATGGAACATCAACTAATTTGTAGCTTCCCAACTGAATGTAATAATG	1075	
QY	1218	TGCTGCTACTTCTACACCAACGAGCCGACTGGTGGCTGGAATCGACACCTGTAC	1277	
Db	1076	TTCTGCTGCTTTTTTGGCATCAAAACAACATGGTTTATACAGCAGGTACTATACATGTCAC	1135	
QY	1278	CTCTGTTAAACAAGCTGACCTCTGGAGCTGAGGCT	1314	
Db	1136	TGAATGTACTAAAAAATAACTTCTGGTCCACAGCT	1172	
RESULT 6				
AF429315/c				
LOCUS				
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.			
ACCESSION	AF429315			
VERSION	AF429315.1			
KEYWORDS	GI:17646244			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 125020)			
AUTHORS	Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S., Ingersoll-Ashworth, R.G., Fleisher, A., Stevanin, G., Brice, A., Potter, N.T., Ross, C.A. and Margolis, R.L. A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2 Nat. Genet. 29 (4), 377-378 (2001)			
JOURNAL	21583737			
MEDLINE	11694876			
PUBMED	2 (bases 1 to 125020)			
REFERENCE	Holmes, S.E., Ingersoll-Ashworth, R.G., Ross, C.A. and Margolis, R.L. Direct Submission			
AUTHORS	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA			
TITLE	Location/Qualifiers			
JOURNAL	1. .125020			
MEDLINE	/organism="Homo sapiens"			
PUBMED	/db_xref="taxon:9606"			
REFERENCE	/chromosome="16"			
AUTHORS	/map="16q24.3; between D16S520 and WI-12410"			
JOURNAL	/note="Isolated from a patient with Huntington's Disease-Like 2 (HDL2)"			
FEATURES	complement(35581. .35746)			
source	/rpt_type=tandem			
	/rpt_unit=ctg			
	complement(<36507. .>36887)			
	/gene="JPH3"			
	/note="JP3"			
	complement(<36507. .>36887)			
	/gene="JPH3"			
	/product="Junctophilin 3"			
	/protein_id="AAL40941.1"			
	/db_xref="GI:17646245"			
	/translation="MSSGRPNFDDGSGYCGWEDKAHGHVCTGPKQGGYTGWSV HGFVLGVYITWPSGNITQGTWAGKRHGIGSKGKWTYKGEWTHFGKRGVRCAG			

Thu Feb 20 11:10:35 2003

BASE COUNT	29056 a	32731 c	30696 g	28283 t	4254 others
ORIGIN	NGAYECTWSNGLQDGYGTETYSDG*				
Query Match	3.6%; Score 50; DB 9; Length 125020;				
Best Local Similarity	10.1%; Pred. No. 0.0056;				
Matches	93;	Conservative	409;	Mismatches	406; Indels 17; Gaps 3;
QY	274	AAGCTGAAGTCTCGCTGGAAACCGCTATCGCTGAGGAGCTACCGACTACGCTGCCTATC	333		
Db	17784	AGKSRGCTGYMKKGGKGSMSKWKGSSTRRRGSAAKSCSYMWGMSMCMSCMSM	17725		
QY	334	ATCACCGAGTGTGAAGCTGTCGATCACTTACACGAGAGAGCTCTCACTTCAAC	393		
Db	17724	AKSYMICYMYRMSYMSYKYKSCMGMSSTSYSCWKMSWGSYKCMKYISGWS	17665		
QY	394	GCTGGAGCTTCTACCTGTACCGCTGTCTGTGAACCGCTGGGAGGAGCTCT	448		
Db	17664	SYSTSMGYSSSTCKKYKCSWSMYKCKTSKYRKRYSYYWGGKRAKKYKACGRRR	17605		
QY	449	CCGCTGGAACGCTGCTACCATCGTGGCTACGTGTAACGTGCTGTCTACCGGACCG	508		
Db	17604	MSYWKCAKWMWSYCCWSYCMYVYKSKCTYKSRGGYWGSKTCYSAGGKSRMY	17545		
QY	509	CTCTGGACGAGGAGTGACCGGACTACGTGGCTCTTACCGAGT-GTGTGAAGTGT	567		
Db	17544	YCMRRSSKSSWSMARSWCNGWAGYRRSKRAGWRRSKGRSTGMKRAKCSK	17485		
QY	568	CGCTGAACCTTCTACTACACGGAACAAACGGAACACCCCTTCAACCCCTGGAAAGTCT	627		
Db	17484	TGSGVSTGRSMKKKGYSKYSRGMKKKTKCYCMKKYKRYKTSWCMYMYKMSWGYKRYK	17425		
QY	628	CAGTGTACCCCTGTCTGTATCAAGCTGTACAGTGGCTCAGCTACCTCGGAAAC	687		
Db	17424	RCCMKKKGCTGYRGMSSKSGYKSMRGRSSYSTSCWKSWMKMYKSKYKRR	17365		
QY	688	GACCTACCATCAGCTGAGTGTACGTGGCTGTCTGAGGGAACCACTCTCTCTGCT	747		
Db	17364	SMRGSMSKWRGVAGRGCTSSSWSTRKRSKCYKSYKKGKGMKMGKMGKRSKYW	17305		
QY	748	GGAGTGAACACTGGGTGGCTAGACACCGAGTGTACCACTGTCTCTCACTTCTAC	807		
Db	17304	SSMKMKRSMKCYSTKSYKSGRRSKGWRSTKSKAKSSMRMAKSKCTYGSSTWSNR	17245		
QY	808	AACAACACGCTCTCACTCAACCCCTGGAAACTCTACCTGCTGCTGCTCTCTCTGCTAAC	867		
Db	17244	RNNR-----MGKTGCNMYRBSRAMNGNAAAGCTTCCCCANTNGGGGAAAA	17196		
QY	868	AAGGACTACGAGCTGAGGCTACCGCTGGAGGAGCTGCTACCTGGCTAAGCAGTGAAC	927		
Db	17196	AGGGGGSASRASCYKGRMSKSYRSGTRRRCKMSKRRSGKSMGMTGRSGGKTSYS	17136		
QY	928	ATCGCTTCTCTGAGGAAACCGCTATCGCTTCTGAGCTACCAACTACGTGATCTCTGCGAG	987		
Db	17136	AKSGRGCTGYCWGKGRKCMSSRKMKMYKSYRERWTCMKCYSMYMYMYCRSMCC	17076		
QY	988	ACCGAGTGTGAAGTGTGCTGCTAATCTTCTACTTCGACGGAACAACTTCCAGGTGA	1047		
Db	17076	MCCKSKCCGYSWGMSSYSYSGYKSWGMKSYMWSYYSKRSTSKAWRSKRGMGTGGRYK	17016		
QY	1048	TCTTCTCGCTGTAAGCTGTCTCTGCTAACAGAGTGCAGGAGCTGTGGCTACCGCTGGA	1107		
Db	17016	GGRSYKGGGYSKGGGSGWGGKGRSGSAGKRSYKMKSSCARYSKMSYKCYSRWM	16956		
QY	1108	GGAACCGCTACCGCTGATCGCTCAGTGTGCTGCTGGAGTGTCTGCTGGAACCGCTGAC	1167		
Db	16956	WMSYIYCWGKWCSSWSSMRKSKGWSAKGMVWDKGSVSTDKSKMBSRSBSKVS	16896		
QY	1168	GACGGAACCACTCTACCTACAGC	1192		
Db	16896	KWSMRGRGYCTTDSMKMYITGC	16871		

RESULT 7
AC073693/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT

AC073693 230372 bp DNA linear HTG 29-JUN-2000
Mus musculus clone RP23-152L22, WORKING DRAFT SEQUENCE, 19
unordered pieces.
AC073693
AC073693.1 GI:8810310
HTG; HTGS_PHASE1; HTGS_DRAFT.
Mus musculus.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
DOE Joint Genome Institute.
1 (bases 1 to 230372)
Sequencing of Mouse
Unpublished
2 (bases 1 to 230372)
DOE Joint Genome Institute.
Direct Submission
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1795526
Center clone name: RPCI-23_152L22

Summary Statistics
Consensus quality: 217356 bases at least Q40
Consensus quality: 223517 bases at least Q30
Consensus quality: 224719 bases at least Q20
Estimated insert size: 258000; adarose-fp estimation
Estimated insert size: 228572; sum-of-contigs estimation
Quality coverage: 7.65 in Q20 bases; agarose-fp estimation
Quality coverage: 8.64 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1087: contig of 1087 bp in length
* 1088 1187: gap of unknown length
* 1188 2715: contig of 1528 bp in length
* 2716 2815: gap of unknown length
* 2816 4158: contig of 1343 bp in length
* 4159 4258: gap of unknown length
* 4259 6473: contig of 2215 bp in length
* 6474 6573: gap of unknown length
* 6574 8982: contig of 2409 bp in length
* 8983 9082: gap of unknown length
* 9083 11186: contig of 2104 bp in length
* 11187 17000: contig of 5714 bp in length
* 17001 17101: gap of unknown length
* 17101 25179: contig of 8079 bp in length
* 25180 25279: gap of unknown length
* 25280 34875: contig of 9596 bp in length
* 34876 34976: gap of unknown length
* 34976 43664: contig of 8689 bp in length
* 43665 53788: gap of unknown length
* 53789 53889: contig of 10024 bp in length
* 53890 65220: contig of 11632 bp in length
* 65221 80284: gap of unknown length
* 80285 80385: contig of 14664 bp in length
* 80386 98449: gap of unknown length
* 98450 98499: contig of 18065 bp in length
* 98500 98549: gap of unknown length
* 98550 98599: contig of 11632 bp in length
* 98600 98649: gap of unknown length
* 98650 98699: contig of 14664 bp in length
* 98700 98749: gap of unknown length
* 98750 98799: contig of 18065 bp in length
* 98800 98849: gap of unknown length
* 98850 98899: contig of 11632 bp in length
* 98900 98949: gap of unknown length
* 98950 98999: contig of 14664 bp in length
* 99000 99049: gap of unknown length
* 99050 99099: contig of 18065 bp in length
* 99100 99149: gap of unknown length
* 99150 99199: contig of 11632 bp in length
* 99200 99249: gap of unknown length
* 99250 99299: contig of 14664 bp in length
* 99300 99349: gap of unknown length
* 99350 99399: contig of 18065 bp in length
* 99400 99449: gap of unknown length
* 99450 99499: contig of 11632 bp in length
* 99500 99549: gap of unknown length
* 99550 99599: contig of 14664 bp in length
* 99600 99649: gap of unknown length
* 99650 99699: contig of 18065 bp in length
* 99700 99749: gap of unknown length
* 99750 99799: contig of 11632 bp in length
* 99800 99849: gap of unknown length
* 99850 99899: contig of 14664 bp in length
* 99900 99949: gap of unknown length
* 99950 99999: contig of 18065 bp in length

```

* 98450 98549: gap of unknown length
* 98550 116262: contig of 17713 bp in length
* 116263 116362: gap of unknown length
* 116363 140706: contig of 24344 bp in length
* 140707 140806: gap of unknown length
* 140807 167609: contig of 26803 bp in length
* 167610 167709: gap of unknown length
* 167710 199398: contig of 31689 bp in length
* 199399 199498: gap of unknown length
* 199499 230372: contig of 30874 bp in length.

FEATURES
    source
        1..230372
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="RP23-152L22"
            /clone_lib="RPCI mouse BAC library 23"
BASE COUNT 61266 a 53740 c 52522 g 61042 t 1802 others
ORIGIN

```

```

Query Match      3.1%; Score 43.4; DB 2; Length 230372;
Best Local Similarity 53.9%; Pred. No. 0.51;
Matches 89; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 166 GCTGCTGCTTCGCTGGAGCTCTACCTGTACCCCTTCCTCAGAGAAGACGCT 225
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212511 GCTGCTGCTCCCTGGCTGCTGCTCCTCCCTGGCTGCTGCTGCTCCTGGCTGCT 212452

QY 226 GGAGCTACGCCAACCTCCTGCTACCGCTAACCTGTTGCCACCGAGTGTACGTTCAAGTGT 285
      ||| || || ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 212451 GCTGCTGCTCCCTGGCTGCTGCTACTGCTCCCTGGATGCTGCTGCCCTGAATGGT 212392

QY 286 CCTGCTGGACCGCTATCGCTGGAGGAGCTACCGACTACGCTGCT 330
      ||||| || ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212391 GCTACTGCTCCCTGGCTGCTGCTGCTGCTCCCTGGATGCTGCT 212347

```

```

RESULT 8
AC093866
LOCUS      AC093866              159764 bp    DNA    linear    PRI 01-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-582C12 from 4, complete sequence.
ACCESSION  AC093866 AC025608
VERSION    AC093866.3 GI:16973777
KEYWORDS   HTG.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Sulston,J.E. and Waterston,R.
            Toward a complete human genome sequence
            Genome Res. 8 (11), 1097-1108 (1998)
            99063792
            9847074
REFERENCE  2 (bases 1 to 159764)
            Pearman,C., Meyer,R. and Doeberber,A.
            The sequence of Homo sapiens BAC clone RP11-582C12
            Unpublished (2001)
REFERENCE  3 (bases 1 to 159764)
            Waterston,R.H.
            Direct Submission
            Submitted (10-SEP-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            4 (bases 1 to 159764)
            Waterston,R.H.
            Direct Submission
            Submitted (17-NOV-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            5 (bases 1 to 159764)
            Waterston,R.H.
            Direct Submission
            Submitted (03-JAN-2002) Genome Sequencing Center, Washington

```

```

University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 159764)
Waterston,R.
Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 17, 2001 this sequence version replaced gi:15987350.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0582C12
Drafting Center: WIBR
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tatenoe,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletier de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACes.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-549C16, 2000 bp overlap; the clone sequenced to the right is RP11-115D19. Actual start of this clone is at base position 123633 of RP11-549C16; actual end is at base position 159764 of RP11-582C12.

Sequence derived from one plasmid subclone, base position 55515 to 55572.

The sequence of AC025608 has been incorporated into AC093866.

FEATURES	source
1..159764	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/chromosome="4"	
/map="4"	
/clone="RP11-582C12"	
/clone_lib="RPCI-11"	
137..299	repeat_region
/rpt_family="ERV1"	
302..653	repeat_region
/rpt_family="ERV1"	
646..735	repeat_region
/rpt_family="ERV1"	
811..11176	repeat_region
/rpt_family="MaLR"	

repeat_region	2191..2508	/rpt_family="MER2_type"
repeat_region	2504..2654	/rpt_family="MER2_type"
repeat_region	2771..3012	/rpt_family="MaLR"
repeat_region	3865..3900	/rpt_family="(CA)n"
repeat_region	4553..4848	/rpt_family="Alu"
repeat_region	5091..5156	/rpt_family="CA-rich"
repeat_region	5387..5515	/rpt_family="Alu"
repeat_region	5562..5671	/rpt_family="L2"
repeat_region	6206..6570	/rpt_family="MaLR"
repeat_region	6680..7334	/rpt_family="L1"
repeat_region	7335..7712	/rpt_family="MaLR"
repeat_region	7713..7960	/rpt_family="L1"
repeat_region	7961..8248	/rpt_family="Alu"
repeat_region	8249..8674	/rpt_family="L1"
repeat_region	8771..8846	/rpt_family="MIR"
repeat_region	8920..8977	/rpt_family="CT-rich"
repeat_region	9169..9212	/rpt_family="AT-rich"
repeat_region	9647..9709	/rpt_family="L2"
repeat_region	10764..11375	/rpt_family="L1"
repeat_region	11377..11733	/rpt_family="MaLR"
repeat_region	11741..12592	/rpt_family="L1"
repeat_region	12593..12815	/rpt_family="MER1_type"
repeat_region	12816..13318	/rpt_family="L1"
repeat_region	13323..13372	/rpt_family="MER1_type"
repeat_region	13830..14371	/rpt_family="ERV1"
repeat_region	14621..14680	/rpt_family="L1"
repeat_region	14884..14816	/rpt_family="Alu"
repeat_region	15382..15437	/rpt_family="MER1_type"
repeat_region	15438..15875	/rpt_family="ERV1"
repeat_region	15876..16041	/rpt_family="MER1_type"
repeat_region	16406..16850	/rpt_family="ERV1"
repeat_region	17833..18128	/rpt_family="MaLR"
repeat_region	18115..18138	/rpt_family="AT-rich"
repeat_region	18134..18238	/rpt_family="Alu"
repeat_region	18217..18239	/rpt_family="(TAAA)n"
repeat_region	19569..19811	/rpt_family="L1"
repeat_region	19914..20213	
repeat_region		/rpt_family="L1"
repeat_region		/rpt_family="AT-rich"
repeat_region		20361..20396
repeat_region		/rpt_family="AT-rich"
repeat_region		20808..20957
repeat_region		/rpt_family="Achobo"
repeat_region		23102..24306
repeat_region		/rpt_family="L1"
repeat_region		24291..24321
repeat_region		/rpt_family="AT-rich"
repeat_region		24311..24706
repeat_region		/rpt_family="L1"
repeat_region		24833..24881
repeat_region		/rpt_family="AT-rich"
repeat_region		24843..25094
repeat_region		/rpt_family="Achobo"
repeat_region		25106..25129
repeat_region		/rpt_family="AT-rich"
Query Match 2.9%; Score 41.4; DB 9; Length 159764;		
Best Local Similarity 52.6%; Pred. No. 1.9;		
Matches 90; Conservative 0; Mismatches 81; Indels 0; Gaps 0;		
QY	82	GAGACCAACCGCTGGACAGGTGGACGACCTGGCAACCCCTGCTAACTGTGTAAGTGT 141
Db	133672	GAAGCCTTCAGAGAGGATGGGTGACTGCTCTGGGTGCCAGGATAGGGGGTTCACAGA 133731
QY	142	CAGAAGAATTTCTACTACAAACCGTGTCTTTCGTGGCGCTGGAGCTTCTACCTGTACC 201
Db	133732	GTAAGGACATCTACTGCTACAACCATGTTCTCTGGAGCTATAAACTAGGAAGTGAACC 133791
QY	202	CCTTGTCTCAGAAAGGACGCTGGAGCTCAGCCTAACCCCTCCCTGCTACC 252
Db	133792	TATAGTCATCAGCAGCTGAATATGAGAGGTGAGAAACAATCTATGGACC 133842
RESULT 9		
AF374493	643 bp	DNA linear PLN 27-JUL-2001
LOCUS	Thalassiosira weissflogii isolate CCMP 1336 clone 10 sexually	
DEFINITION	Induced protein SIG 1 gene, partial cds.	
ACCESSION	AF374493	
VERSION	AF374493.1	GI:143336349
KEYWORDS	Thalassiosira weissflogii.	
SOURCE	Thalassiosira weissflogii.	
REFERENCE	1 (bases 1 to 643)	
AUTHORS	Armbrust, E.V. and Galindo, H.M.	
TITLE	Rapid evolution of a sexual reproduction gene in centric diatoms of	
JOURNAL	the genus Thalassiosira	
MEDLINE	Appl. Environ. Microbiol. 67 (8), 3501-3513 (2001)	
PUBMED	21365169	
REFERENCE	2 (bases 1 to 643)	
AUTHORS	Armbrust, E.V. and Galindo, H.G.	
TITLE	Direct Submission	
JOURNAL	Submitted (27-APR-2001) Oceanography, University of Washington, Box	
FEATURES	357940, Seattle, WA 98195-7940, USA	
source	Location/Qualifiers	
	1..643	/organism="Thalassiosira weissflogii"
		/isolate="CCMP 1336"
		/db_xref="taxon:67004"
		/clone="10"
		/country="USA; Long Island, NY"
		join(<1..56,141..>643)
		/product="sexually induced protein SIG 1"
		join(<1..56,141..>643)
		/codon_start=2
mrna		
CDS		

[illegible]

LOCUS	HSCOLPDGF	16814 bp	DNA	linear	PRI 12-MAR-1997
DEFINITION	H.sapiens DNA sequence of COL1A1 gene fused with intron 1 of PDGFB gene.				
ACCESSION	X98705				
VERSION	X98705.1	GI:1888408			
KEYWORDS	COL1A1 gene; collagen alpha 1 type I; PDGFB gene; platelet-derived growth factor beta; translocation breakpoint.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Simon,M.P., Pedetour,F., Sirvent,N., Grosgeorge,J., Minioletti,F., Colindre,J.M., Terrier-Lacombe,M.J., Mandahl,N., Craver,R.D., Blin,N., Sozzi,G., Turc-Carel,C., O'Brien,K.P., Kedra,D., Franss,I., Guilbaud,C. and Dumanski,J.P.				
TITLE	Deregulation of the platelet-derived growth factor B-chain gene via fusion with collagen gene COL1A1 in dermatofibrosarcoma protuberans and giant-cell fibroblastoma				
JOURNAL	Nat. Genet. 15 (1), 95-98 (1997)				
MEDLINE	97141927				
PUBMED	8988177				
REFERENCE	2 (bases 1 to 16814)				
AUTHORS	O'Brien,K.P.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-JUN-1996) K.P. O'Brien, Dept. of Molecular Medicine, Clinical Genetics Unit, Karolinska Hospital, Building L-6, S-171 79 Stockholm, SWEDEN				
COMMENT	Related sequences: J03559, M10627, X07884, M20789 and K01228.				
FEATURES	Location/Qualifiers				
source	1..16814				
	/organism="Homo sapiens"				
	/isolate="cloned in cosmid pWE15 from tumor DNA from dermatofibrosarcoma protuberans (tumor TNM-1241-90)"				
	/db_xref="taxon:9606"				
	/clone_lib="cosmid D4"				
	/note="chimaeric cosmid clone spanning a translocation break between COL1A1 gene fused with intron 1 of PDGFB gene"				
source	1..15822				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="17"				
	/map="q21.3-q22"				
source	15831..16814				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="22"				
	/map="q13"				
TATA_signal	2456..2460				
gene	2484..15561				
mRNA	Join(2484..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)				
	/product="collagen type I alpha 1"				
exon	2484..2705				
	/gene="COL1A1"				
	/number=1				
CDS	Join(2603..2705,4168..4362,4501..4535,4638..4673,4764..4865,5588..5659,5887..5931,6090..6143,6306..6359,6858..6911,7028..7081,7421..7474,7563..7607,7724..7777,7892..7936,8115..8168,8426..8524,8612..8656,8760..8858,8990..9043,9261..9368,9464..9517,9643..9741,9907..9960,10049..10147,11045..11098,11242..11295,11399..11452,11564..11617,12068..12112,12206..12304,12602..12709,13168..13275,13496..13549,13712..13765,13983..14090,14179..14232,14359..14412,14553..14714,14816..14923,15081..15188,15296..15349,15454..>15561)</				

exon	6306..6359 /gene="COL1A1" /number=9
intron	6360..6857 /gene="COL1A1" /number=9
exon	6858..6911 /gene="COL1A1" /number=10
intron	6912..7027 /gene="COL1A1" /number=10
exon	7028..7081 /gene="COL1A1" /number=11
intron	7082..7420 /gene="COL1A1" /number=11
exon	7421..7474 /gene="COL1A1" /number=12
intron	7475..7562 /gene="COL1A1" /number=12
exon	7563..7607 /gene="COL1A1" /number=13
intron	7608..7723 /gene="COL1A1" /number=13
exon	7724..7777 /gene="COL1A1" /number=14
intron	7778..7891 /gene="COL1A1" /number=14
exon	7892..7936 /gene="COL1A1" /number=15
intron	7937..8114 /gene="COL1A1" /number=15
exon	8115..8168 /gene="COL1A1" /number=16
intron	8169..8425 /gene="COL1A1" /number=16
exon	8426..8524 /gene="COL1A1" /number=17
intron	8525..8611 /gene="COL1A1" /number=17

Query Match	2.78;	Score 38.6;	DB 9;	Length 16814;
Best Local Similarity	46.7%;	Prod. No. 11;		
Matches 122;	Conservative	0;	Mismatches 139;	Indels 0;
Gaps	0;			
QY	440	GAGCTCTGACCGCTGGAAACGCTGCTACCATCTGGTTCAGTGTACGTGGCTTGTTCCTA	499	
Db	8568	GAGGTTAGAAAGTGGCAAAAGGACACTGACTGGGGGACACTTACGACGGGCCACGACGA	8509	
QY	500	CGGGAACCGCTCTGACGACGAGTGAACACCGACTACGTGGCGCTTTTCCACGAGCTGTG	559	
Db	8508	CCAGCAGGGCCAGGGGGCCAGGCTTCACAGCACACCTTGGGACCTTCAGAGCCTCG	8449	
QY	560	TGAAGTGTGCGCTGAACCTTCTACTAACCGGAAACACGGAAACACCCCTTTTCAACCCCTG	619	
Db	8448	GGCCCTTGGGACACGAGCTTCAACCTTGAATCAGAGAAAGGACATATCAGAAGGCCACCTG	8389	
QY	620	GAAAGTCTCAGTGTACCCCTTCTCTCTGCTATCAAGCCCTGTCTAAAGCTGGCTTCAGGCTACCC	679	

Df 8398 GGAACCCACACTTGCCTCTCGGATGGCAGGAGGAAGGACACAGGACAGGAGCA 8329

Qy 680 TGGAAACGACGCTACCATCA 700
| | | | | | | | | |
Df 8328 ATGATCAGGACTCTAAGATCA 8308

RESULT 14
AP003600/c
LOCUS
DEFINITION Nostoc sp. PCC 7120 plasmid pCC7120alpha DNA, complete genome,
section 1/2.
ACCESSION AP003600 BA000020
VERSION AP003600.1 GI:17135539
KEYWORDS
SOURCE Nostoc sp. PCC 7120 plasmid:pCC7120alpha DNA.
ORGANISM Nostoc sp. PCC 7120
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

REFERENCE
AUTHORS Kaneko,T., Nakamura,Y., Wolk,C.P., Kuritz,T., Sasamoto,S.,
Watanabe,A., Iriguchi,M., Ishikawa,A., Kawashima,K., Kimura,T.,
Kishida,Y., Kohara,M., Matsumoto,M., Matsuno,A., Muraki,A.,
Nakazaki,N., Shimpo,S., Sugimoto,M., Takazawa,M., Yamada,M.,
Yasuda,M. and Tabata,S.
Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120
DNA Res. 8 (5), 205-213 (2001),
21595285
2 (bases 1 to 341950)
Kaneko,T.
Direct Submission
Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research; Yana
1532-3, Kisarazu, Chiba 292-0812, Japan
(E-mail:kaneko@kazusa.or.jp/cyanobase/
URL:http://www.kazusa.or.jp/cyanobase/
Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)

FEATURES
source
1. 341950
/organism="Nostoc sp. PCC 7120"
/db_xref="taxon:103650"
/plasmid="pcc7120alpha"
/note="synonym:Anabaena sp. PCC7120"
complement(165..548)
/gene="all7002"
complement(165..548)
/gene="all7002"
/note="ORF_ID:all7002"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB78086.1"
/db_xref="GI:17135540"
/translation="MVRKSYPDTLDMWEILSPLIIPAKEGGHPRTDMREICNALY
YHLTKGCONMLPGDFPPSTVSVYRKWKRGVMEKLHNTLRGVRAKLKGSTQPSPA
IADSSSVKTAERGMCIALMELKR"
1123..1719
/gene="alr7003"
1123..1719
/gene="alr7003"
/note="ORF_ID:alr7003"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="BAB78087.1"
/db_xref="GI:17135541"
/translation="MEKVQNFETFKVHEVTANNLSIAEANKSPVVPRPQNTSLX
RKVQSAKAARRDDIHARVMSLRISGLSQAIQAOLSGVAKNITVNYLRSSFTRR
SDKGLSLNPYHDYLISRNSNGHSOKLFEEIRAYGTGSYATVSREFTRYIKTVPSA
EPAGSRKKNASLRRVSSSHRLPTPSRVTLALRRPELI"
complement(1734..2477)
/gene="all7004"
complement(1734..2477)
/gene="all7004"
complement(1734..2477)
/gene="all7004"

